

EARTHQUAKE PROBABILITIES IN THE KUMAUN-GARHWAL HIMALAYA FROM STOCHASTIC MODELING OF MODERATE-TO-LARGE SEISMIC EVENTS

S. PASARI¹ & S. K. ARORA²

¹Birla Institute of Technology and Science, Pilani, Rajasthan, India

²Ex-Head, Seismology Division, Bhabha Atomic Research Centre, Mumbai, Maharashtra, India

ABSTRACT

Empirical statistical modeling of earthquake interevent time-gaps has over the years generated enormous interest to geoscientists and other professionals, for a variety of applications relating to mitigation of earthquake risk. In this study, we estimate probabilities of earthquake recurrence in the Kumaun-Garhwal region of the central Himalayan belt, using a set of internationally catalogued 44 moderate-to-large ($5 \leq m_b \leq 6.3$) independent earthquakes, that occurred there during a period of 60 years (1958-2017). We subjected the interevent time intervals of these earthquakes to detailed stochastic processing, to examine the efficacy of each of the four different probability distribution models, viz. exponential, gamma, lognormal and Weibull, and found that, the lognormal model best represents the observed data in this case. The statistical inferences drawn are based essentially on two goodness-of-fit tests: (i) Maximum Likelihood Estimation (MLE) with an improvement over the Akaike Information Criterion (AIC) and (ii) the non-parametric Kolmogorov-Smirnov (K-S) test. In addition, the surrogate Fisher Information Matrix (FIM) approach is utilized for uncertainty estimation. As a measure of seismic hazard, we determined the earthquake potential in the study region, in terms of conditional probabilities of future earthquakes, for varying elapsed times of assumed seismic quiescence, since the last event detected on 06 February 2017 at the time-complete catalog.

KEYWORDS: Earthquake, Kumaun-Garhwal Himalaya, Probability

INTRODUCTION

Earthquakes occur as a consequence of global plate tectonic motions leading to deformation, that causes progressive build-up of strain in the upper brittle part of the earth's crust. The elastic stress steadily accumulating in the lower ductile part of the crust also gets continually transferred to the upper crust part which is eventually stressed to a critical limit resulting in fault rupture. The stress drop after relaxation resets stage for the next cycle of earthquake preparation, and so on. Such interactions among different spatio-temporal systems, govern the complex nonlinear Geodynamics of earthquake sequences. Despite this reality, one may also consider earthquakes as a simple point process in space and time, by ignoring its temporal nature due to limited event-time duration and spatial nature, due to a finite extent of a rupture zone (Utsu 1984; Anagnos and Kiremidjian 1988). Besides, earthquake processes in seismically active zones are observed to be partially stochastic (Shen et al. 2007; Vere-Jones 2009). Thus, it is important to study various statistical properties of earthquake sequences in a given region towards assessment of earthquake hazard in terms of conditional probability of future earthquake events in that region. Such a study serves a useful purpose in a wide variety of applications, including engineering designs, city planning, archaeological preservations and, indeed, estimation of

earthquake risk to structures on the ground and human population, based on their vulnerability quotient.

The present study area lying within the grid bounded by 29-31°N and 79-82°E (Figure 1) Encloses the Kumaun-Garhwal region of the Central Himalaya. This particular region is seismotectonically quite active as it is traversed by a number of major fault zones that keep reactivating time and again causing moderate to large earthquakes. A south, south-verging megathrust fault forming the Main Central Thrust (MCT) with its several imbricate branching faults separate this region from the Main Boundary Thrust (MBT). The lesser Himalayan belt between the MBT and MCT also has sub-units demarcated by several major thrusts. Strongly folded and imbricated weakly metamorphosed sedimentary series of rocks, crystalline sequence of medium-to-high grade metamorphic rocks of Ordovician and early Miocene age, and fossiliferous sediments of Upper Proterozoic to Middle Eocene age constitute the geotectonic setting of the study region (Valdiya 1998; Yin 2006). The root of several nappes, klippen, duns, and lakes can be observed in this sub-unit of the Himalayan orogen (Valdiya 1998).

Physical and deterministic modeling of earthquake sequences from geological, paleoseismological or geodetic data are, no doubt, valuable. However, these approaches necessitate to gather field observations intensively over prolonged periods of time and thus are often expensive, cumbersome and time consuming. Therefore, as an alternative approach, it is motivating to study statistically the temporal characteristics of earthquake intervened times leading to an assessment of seismic hazard.

In the present work, we concentrate on the empirical statistical distribution of time intervals between successive earthquakes in the Kumaun-Garhwal Himalaya. We examine the applicability of different models, viz. Exponential, gamma, lognormal, and Weibull. Although, each of these distributions has notable advantages in earthquake modeling (Utsu 1984), we identify a model that best represents the observed data of earthquakes in that region. It may be noted that we do not need to take into account faulting parameters including source mechanism, hypocentral depth, source volume, etc.. Our main aim is to determine conditional probability of future earthquakes in the study region based on the best-fit probability model. Such conditional probability estimates also serve to assess seismic hazard in the region.

REGIONAL EARTHQUAKE DATA

The present study utilizes real, complete and homogeneous data set of 44 moderate-to-large earthquakes in the body-wave magnitude range ($5 \leq m_b \leq 6.3$) that occurred in the study region over a period of 60 years (1958–2017). All of these earthquakes are shallow-focus with their sources within the crust at depths mostly not exceeding about 50 km (Figure 1). These data have been adopted from the earthquake catalog of the International Seismological Centre (ISC) after a careful scrutiny by picking up only independent events, dropping out the dependent ones such as foreshocks, aftershocks and seismic swarms or clusters. The upper limit of magnitude (m_b 6.3) has been consistent with the largest value registered over the data period since 1958, while the lower cutoff magnitude (m_b 5.0) is set by us taking into account that earthquakes smaller than this size are of no consequence worth considering from the hazard point of view. The epicentral locations, focal depths, magnitudes, and times of occurrence of these earthquakes are summarized in Table 1. The largest event (m_b 6.3) in the compiled list severely jolted the Chamoli district of Uttarakhand on March 28, 1999, whereas the last listed event (m_b 5.3) that occurred near Pipalkoti in Uttarakhand on February 06, 2017 produced moderate ground motion. However, these earthquakes generated considerable discussion on engineering consequences as experienced during the long-debated Tehri Dam construction activity in the Garhwal Himalaya (Kayal 2001; Bisht 2008).

The uncontaminated data set constituted by earthquake events selected in the present study ensures that the random samples of earthquake intervened times are truly independent and identically distributed (hereafter referred to as IID type events)). For obtaining such an IID data set, a window-based spatio-temporal type filtering is adopted. This approach aims to appropriately identify foreshocks and aftershocks, and exclude them from the list, being disqualified for the present analysis. The criterion used relies on the following conditions, according to which a dependent event is one that (a) it lies on the same fault which the main (parent) earthquake ruptured and its source is not far away (typically, within about 50 km) from that of the parent event, (b) it precedes or succeeds the main event, within a small time marking temporary quiescence, and (c) it is necessarily of a lower magnitude, quite often much lower, as compared to the main event.

In our data set, we need to homogenize magnitudes. We chose to retain the body-wave magnitude (m_b) as standard, since a large majority of the events in the original ISC catalog is assigned such a magnitude. We converted magnitudes of the remaining events, reported with either local magnitude (M_L) or surface-wave magnitude (M_s), making use of the following well known empirical relations (Gutenberg and Richter 1956; Abe 1981; Kanamori 1983; Scordilis 2006).

$$m_B = 1.5m_b - 2.2 \quad m_B = 0.63M_s + 2.5 \quad M_s = 1.27(M_L - 1) - 0.016M_L^2 \quad (1)$$

The magnitude m_b or m_B follows from the short-period (about 1 sec) initial P-wave amplitude, which was introduced by Gutenberg (1945a, 1945b) after the installation of the World-Wide Standardized Seismograph Network (WWSSN) to routinely record the short-period vertical-component seismograms. The surface-wave magnitude M_s is usually determined from the maximum amplitude of Rayleigh type long-period seismic surface waves with a period of about 20 Sec (Gutenberg 1945a), whereas the local magnitude M_L is determined from the maximum signal amplitude recorded on a Wood-Anderson seismogram with predominant period usually within 0.1–3.0 Sec (Richter 1935).

In stochastic modeling and analysis of earthquake occurrences, as in the present study, completeness of the print catalog of reported earthquakes above a certain magnitude value is of utmost importance as a prerequisite to detailed processing. In this regard, we performed a magnitude-frequency based visual cumulative method (VCM) test (Mulargia and Tinti 1985) on the original ISC catalog of earthquakes covering the 60-year period (1958-2017), in the operative magnitude range ($5 \leq m_b \leq 6.3$). A linear fit in the least-squares sense is obtained from the scatter plot between time (in years) and the cumulative number of earthquake events in the chosen magnitude range. A catalog is designed to be ‘time- complete’ if the trend of the observed data stabilizes to approximately a straight line (Mulargia and Tinti 1985). In other words, it signifies that the earthquake rates and moment releases are ultimately steady over the time periods considered (Mulargia and Tinti 1985). Of course, we are aware that, at times, the possibility of extended aftershock durations in low-strain intercontinental regions (e.g., Stein and Liu 2009) and the possibility of substantial variations in seismic activity (e.g., Page and Felzer 2015) may raise some pertinent questions about the typical linear assumption in respect of the catalog completeness. However, in the present application using the filtered data set, it is observed that the graph between time and the cumulative event count has a linear relationship with R-square value at 0.94 (Figure 2), which evidently satisfies the catalog completeness requirement very well.

METHODOLOGY

Stochastic modeling with regard to earthquake recurrence in a given region generally involves a three-step process. The first step highlights typical assumptions and model descriptions; the second step deals with the model parameter estimation; the third step aims at the model validation from several goodness-of-fit tests. Once the best-fit probability model is identified for the given data set in a region, seismic hazard in terms of estimated conditional probability of earthquake occurrence in that region can be easily determined.

In the present work, four important models of continuous probability distributions, namely exponential, gamma, lognormal and Weibull have been considered. The probability density functions of these distributions along with the respective model parameters, their domains and specific roles are provided in Table 2 (Johnson et al. 1995; Murthy et al. 2004). It is seen that the exponential distribution is controlled by a scale parameter alone, whereas the gamma, lognormal and Weibull distributions are controlled by a scale parameter (log-scale for lognormal) as well as a shape parameter. The scale parameter determines the spread of the distribution, while the shape parameter determines the shape or appearance of the distribution. Between these two controlling parameters, the shape parameter is more important in modeling as it depicts the monotone nature of the hazard function (Johnson et al. 1995). The exponential distribution that does not have any shape parameter always provides only one type of appearance of its density function; it starts at the level of time $t = 0$, and then monotonically decreases exponentially, always convex and stretched to the right as α increases. Moreover, the hazard function of an exponential distribution is constant, signifying that the units (e.g., intervened time gaps) do not degrade with time. Although, this peculiar behavior of exponential distribution seems to contradict in a way the physical assumption in elastic rebound theory, it is considered to be the fundamental distribution in modeling earthquake sequences due mainly to the fact that the number of earthquakes in a region follows a Poisson process. Thus, theoretically, it seems appropriate to argue that earthquake intervened time intervals, which must follow an exponential distribution (see, for more details, Hogg et al. (2005)).

Now, with the known density function $f(t)$ of a positive random variable T , it is straightforward to obtain unique function, hazard distribution function $F(t)$, survival action $S(t)$, hazard function $h(t)$, and reverse hazard function $r(t)$ as follows: $F(t) = \int_0^t f(u) du$, $S(t) = 1 - F(t)$, $h(t) = \frac{f(t)}{S(t)}$, and $r(t) = \frac{f(t)}{F(t)}$. In the present context, the random variable T essentially denotes the earthquake interevent times (time interval between two successive events) providing a random sample $\{T_1, T_2, \dots, T_{43}\}$ of size 43, the total number of selected earthquakes in the present case being 44 (Table 1).

In order to determine the conditional probability of an earthquake above a certain magnitude to occur in an elapsed time, we need to introduce another random variable V , corresponding to a waiting time v . This conditional probability, in the time interval $(\tau, \tau + v)$, given that no earthquake in the same magnitude range has occurred in τ years since the last event in a data set, can be expressed as:

$$P(V \leq \tau + v | V \geq \tau) = \frac{F(\tau + v) - F(\tau)}{1 - F(\tau)} \quad (v > 0) \quad (2)$$

For estimating the scale and shape parameters of the four different probability distributions as mentioned above, the Maximum Likelihood Estimation (MLE) method is adopted not only because of its flexibility and wide applicability, but also for its ability to provide the uncertainty (asymptotic) measure in the estimates. In addition, a Fisher Information Matrix (FIM) based approach is employed as a surrogate tool to examine the uncertainty in the parameter estimates in terms of asymptotic variances and 95% two-sided confidence bounds in the following manner (Hogg et al. 2005).

Let $I_{p \times p}(\theta)$ be the information matrix where $\theta = (\theta_1, \theta_2, \dots, \theta_p)$, for some integer p , denotes the vector of parameters.

Then, $I_{p \times p}(\theta)$ is given by

$$I_{p \times p}(\theta) = (I_{ij}(\theta))_{i,j=1,2,\dots,p} = E \left(- \frac{\partial^2 \ln f(T; \theta)}{\partial \theta_i \partial \theta_j} \right)_{i,j=1,2,\dots,p} = \frac{1}{n} E \left(- \frac{\partial^2 L(T; \theta)}{\partial \theta_i \partial \theta_j} \right)_{i,j=1,2,\dots,p} \quad (3)$$

Where E is the expectation operator and $L(T; \theta)$ is the log-likelihood function of n sample data points $\{t_1, t_2, t_3, \dots, t_n\}$. The FIM is a symmetric and positive semi-definite matrix that provides a measure of the amount of information an observed random sample carries about the unknown population parameter θ . This matrix is often combined with the Cramer-Rao lower bound theorem (Hogg et al. 2005) to estimate asymptotically the variance-covariance matrix $\Sigma_{\hat{\theta}}$ of the estimated parameters $(\hat{\theta})$ as $\Sigma_{\hat{\theta}} \geq [nI(\hat{\theta})]^{-1}$; $\hat{\theta}$ is the maximum likelihood estimate of θ . The $(1-\delta)\%$ two-sided asymptotic confidence bounds on the estimated parameters readily follow from the following inequality:

$$\hat{\theta} - z_{\delta/2} \sqrt{[V_{ii}(\hat{\theta})]_{i=1,2,\dots,p}} < \theta < \hat{\theta} + z_{\delta/2} \sqrt{[V_{ii}(\hat{\theta})]_{i=1,2,\dots,p}} \quad (4)$$

Where $[V_{ii}(\hat{\theta})]_{i=1,2,\dots,p}$ is the vector of diagonal entries in the variance-covariance matrix, and $z_{\delta/2}$ is the critical

value corresponding to a significance level of $\delta/2$ on the standard normal distribution (Hogg et al. 2005).

DATA PROCESSING AND RESULTS

The FIMs of four studied models of probability distributions are presented in Table 2, and for these models, the estimated values of the scale and shape parameters with the uncertainties in their estimates are given in Table 3. It is observed that the estimated shape parameters in both Weibull and gamma distributions are less than 1, which suggests that the seismic hazard function (i.e. The failure rate due to earthquakes) provided by these two probability models is monotonically decreasing in the study region. Furthermore, compared with the exponential model, the Weibull model shows a ‘heavy-tailed’ behavior (the ‘tail’ is relatively thicker) in respect of the earthquake data set in this region.

We now turn to make an assessment as to which of the four probability distribution models best represents the present data set of selected earthquakes. In this regard, we invoke two popular selection tests: the maximum likelihood test with an improvement upon Akaike Information Criterion (AIC) and the non-parametric Kolmogorov-Smirnov (K-S) test. The maximum likelihood test actually utilizes likelihood (or, log-likelihood) values to prioritize different competitive

models. Also, it assumes that the number of parameters in all the candidate distributions is the same. In order to do away with such a presumption, the Akaike criterion (AIC) is in modification expressed as $AIC = 2k - 2\ln L$, where k represents the number of parameters in a model and $\ln L$ denotes the log-likelihood value (Pasari and Dikshit 2014). From such an improved formulation of the AIC, it is clear that it is able to account for the trade-off between a superior fit from number of parameters in a model and the associated model complexity (Hogg et al. 2005). The log-likelihood and AIC values for each of the competing models are given in Table 4.

On the other hand, in the non-parametric K-S test, we first construct the empirical distribution function H_n for n independent (IID type) random variables T_1, T_2, \dots, T_n as

$$H_n(t) = \frac{1}{n} \sum_{i=1}^n I_{T_i \leq t} \quad (5)$$

Where $I_{T_i \leq t}$ is the indicator function, which equals 1 if $T_i \leq t$, and 0 otherwise. This renders $H_n(t)$ a step function. Supposing that we have two competing models F and, the corresponding K-S distances are calculated as

$$\begin{aligned} D_1 &= \sup_{-\infty < t < \infty} |H_n(t) - F(t)| \\ D_2 &= \sup_{-\infty < t < \infty} |H_n(t) - G(t)| \end{aligned} \quad (6)$$

In the above expression (6), \sup_t denotes the supremum of a set of distances. If, we choose model F , otherwise we choose the model. The K-S distance of each of the candidate models is included in Table 4, and their relative closeness (separation) to empirical distribution function is graphically represented in Figure 3.

From a scrutiny of the numerical figures available in Table 4, it turns out that the AIC value of the lognormal probability distribution model stands out minimum (110.97) among those of the other models. This implies that the lognormal distribution is the most optimum fit, considering the trade-off between the model complexity and the model fit, at the interevent time intervals of the present data set of earthquakes in the study region. The K-S test also favours the lognormal distribution, being the 'closest' (smallest separation) to the empirical distribution function of the observed data (Figure 3). Therefore, it is well justified to apply the lognormal probability model to determine the conditional probability of future earthquakes in the study region.

We have consolidated in Table 5 the estimated conditional probabilities of occurrence of an earthquake in the magnitude range 5 to 6.3 in the study region within the next 1 to 5 years, assuming elapsed times (absence of earthquake) over similar time-intervals since the last detected event on 06 February 2017 in the observed data set. These results are illustrated by a family of curves plotted in Figure 4 depicting conditional probabilities of future earthquakes in the coming years in the study region. These curves also serve to infer seismic hazard and the associated risk to life and property in the region.

DISCUSSIONS AND CONCLUSIONS

In recent years, the probability distributions of earthquake intervened times have been regularly used to deduce

important information such as the recurrence pattern and mean waiting time for future earthquakes for various scientific and engineering applications including smart city planning, seismic hazard assessment and risk mitigation, and nationwide earthquake insurance program (Pasari and Dikshit 2014). The best-fit probability model, given a set of systematically catalogued earthquakes, can not only determine future seismicity in active regions, but also be employed as a secondary tool to realize the physical mechanism of earthquake preparation in those regions (Hsu et al. 2009). At times, the probability model formulation is further simplified by reducing the complex dynamics of the 'earthquake machine' to conform to a point process (Utsu 1984). This in turn implies that the complexities involved in the physics of rupture initiation, propagation and stopping are ignored, and the earthquakes are considered to be 'events' characterized only by spatio-temporal occurrence pattern (time, location and magnitude). In this regard, the present study by examining four different probability distributions in the Kumaun-Garhwal region of the Central Himalaya and fixing one that is most optimum projects future seismicity in that region reasonably well.

Quite often, it has been argued that continuous or campaign mode geodetic observations that enable mapping of elastic strain rates through measurements of displacement fields is a more realistic way than the probabilistic approach to earthquake hazard assessment in a seismically active region (Shen et al. 2007). The geodetic data reveal tectonic plate kinematics and the associated Geodynamics, whereas the probabilistic earthquake modeling addresses the stochastic repeating nature of the earthquake process. In order to lend credence to the latter approach, many research groups routinely collect data from networks of GPS (Global Positioning System) to decipher horizontal deformations and strain rates to integrate them with regional historical earthquake data towards an improved understanding of earthquake recurrence processes (England and Molnar 1997; Qin et al. 2002; Pancha et al. 2006; Shen et al. 2007; Hsu et al. 2009). Thus, we are inclined to believe that empirical stochastic modeling can be gainfully combined with geological, paleoseismological and geodetic investigations to address many open questions on earthquake triggering processes in the Himalayan orogen.

For the present, it turns out that the lognormal probability distribution best represents the interevent times through the internationally catalogued earthquakes of moderate to large magnitude over the last six decades in the Kumaun-Garhwal region. The estimated conditional probabilities of future earthquakes in that region, for varying elapsed times of assumed seismic quiescence since the last known event on 06 February 2017, point to moderately large seismic hazard in the central Himalayan domain.

REFERENCES

1. Abe K (1981). Magnitudes of large shallow earthquakes from 1904 to 1980. *Phys Earth Planet Int* 27:72–92
2. Anagnos T, Kiremidjian AS (1988). A review of earthquake occurrence models for seismic hazard analysis. *Prob Engg Mech* 3 (1): 1–11
3. Bisht TC (2008). *Displacement, Resettlement and Everyday Life: An Ethnographic Study of People Displaced by Tehri Dam in India*. Publisher: La Trobe University, p 464
4. England P, Molnar P (1997). The field of crustal velocity in Asia calculated from Quaternary rates of slips on faults. *Geophys J Int* 130:551–582

5. Gutenberg B (1945a). Amplitudes of surface waves and magnitudes of shallow earthquakes. *Bull Seism Soc Am* 35:3–12
6. Gutenberg B (1945b). Amplitudes of P, PP, and S and magnitudes of shallow earthquakes. *Bull Seism Soc Am* 35:57–69
7. Gutenberg B, Richter CF (1956). Magnitude and energy of earthquakes. *Ann Geofis* 9:1–15
8. Hogg RV, Mckean JW, Craig AT (2005). *Introduction to mathematical statistics*. 6th Ed, PRC Press, p718
9. Hsu YJ, Yu SB, Simons M, Kuo LC, Chen HY (2009). Interseismic crustal deformation in the Taiwan plate boundary zone revealed by GPS observations, seismicity, and earthquake focal mechanisms. *Tectonophys* 479:4–18
10. Johnson NL, Kotz S, Balakrishnan N (1995). *Continuous univariate distributions*. Vol 2, 2nd Ed, Wiley, New York. p752
11. Kanamori (1983). Magnitude scale and quantification of earthquakes. In: Duda SJ and Aki H (Editors), *Quantification of Earthquakes*, *Tectonophys* 93:185–199
12. Kayal JR (2001). Microearthquake activity in some parts of the Himalaya and the tectonic model. *Tectonophys* 339 (3-4): 331–351. doi 10.1016/S0040-1951 (01)00129-9
13. Mulargia F, Tinti S (1985). A seismic sampling area defined from incomplete catalogs: an application to the Italian territory. *Phys Earth Planetary Sci* 40 (4), 273–300
14. Murthy DNP, Xie M, Jiang R (2004). *Weibull models*. John Wiley and Sons, New Jersey, 1st Ed. p383
15. Page M, Felzer K (2015). Southern San Andreas fault seismicity is consistent with the Gutenberg-Richter Magnitude-Frequency distribution. *Bull Seismol Soc Am* 105 (4); doi: 10.1785/0120140340
16. Pancha A, Anderson JG, Kreemer C (2006). Comparison of seismic and geodetic scalar moment rates across the Basin and Range province. *Bull Seism Soc Am* 96 (1): 11–32
17. Pasari S, Dikshit O (2014) Impact of three-parameter Weibull models in probabilistic assessment of earthquake hazards, *Pure Appl Geophys* 171 (7): 1251–1281
18. Qin C, Papazachos C, Papadimitriou E (2002). Velocity field for crustal deformation in China derived from seismic moment tensor summation of earthquakes. *Tectonophys* 359:29–46
19. Richter CF (1935). An instrumental earthquake magnitude scale. *Bull Seism Soc Am* 25:1–32
20. Scordilis EM (2006). Empirical global relations converting M_s and m_b to moment magnitude. *J Seism.* 10:225–236
21. Shen ZK, Jackson DD, Kagan YY (2007). Implications of geodetic strain rate for future earthquakes, with a five-year forecast of M5 earthquakes in southern California. *Seism Res Lett* 78 (1): 116-120
22. Stein S, Liu M (2009). Long aftershock sequences within continents and implications for earthquake hazard assessment. *Nature* 462:87–89

23. Utsu T (1984). Estimation of parameters for recurrence models of earthquakes. Bull Earthq Res Inst, Univ Tokyo 59:53–66
24. Valdiya KS (1998). Dynamic Himalaya, Univ Press, Hyderabad. p178
25. Vere-Jones D (2009). Some models and procedures for space-time point processes. Environ Ecol Stat 16:173–195
26. Yin A (2006). Cenozoic tectonic evolution of the Himalayan orogen as constrained by along-strike variation of structural geometry, exhumation history, and foreland sedimentation. Earth Sci Rev 76 (1-2): 1–31

APPENDICES

Table 1: List of Selected Independent Earthquakes ($5 \leq m_b \leq 6.3$) in the Study Region

S.No.	Date			Location		Source Depth	Magnitude (m_b)
	Year	Month	Day	Latitude ($^{\circ}$ N)	Longitude ($^{\circ}$ E)	(km)	(*)
1	1958	12	28	29.926	79.900	15	5.6 (M_S 5.9)
2	1964	9	26	29.960	80.460	50	5.9
3	1965	5	13	29.620	80.190	75	5.1
4	1966	6	27	29.620	80.830	33	6.0
5	1968	1	5	30.410	79.250	07	5.0
6	1968	5	31	29.910	79.920	33	5.0
7	1969	3	3	30.040	79.840	18	5.1
8	1969	6	22	30.500	79.400	15	5.3
9	1976	5	10	29.327	81.458	22.2	5.2
10	1976	9	29	29.502	81.508	33	5.0
11	1978	1	7	30.513	79.404	33	5.1
12	1979	5	20	29.932	80.270	15.8	5.7
13	1980	7	29	29.629	81.091	23.3	6.1
14	1981	3	06	29.799	80.664	23.6	5.1
15	1984	2	19	29.843	80.544	21	5.1
16	1984	5	18	29.520	81.793	x	5.6
17	1990	9	21	29.985	79.907	18.7	5.1
18	1991	12	9	29.512	81.611	02.9	5.6
19	1996	3	26	30.692	79.103	41.8	5.3
20	1997	1	5	29.874	80.565	24.9	5.4
21	1999	3	28	30.511	79.421	22.9	6.3
22	2001	11	27	29.691	81.716	22.6	5.5
23	2002	6	4	30.566	81.420	10	5.4
24	2003	4	4	30.086	80.040	27.7	5.1 (M_L 5.1)
25	2003	5	27	30.556	79.337	28.9	5.0
26	2004	10	27	30.848	81.189	x	5.2 (M_S 5.0)
27	2005	9	5	30.454	79.247	48.3	5.1 (M_L 5.1)
28	2005	10	25	30.127	81.111	41.5	5.0
29	2005	12	24	30.515	79.250	36.9	5.2
30	2006	2	1	30.318	80.388	09.5	5.2 (M_L 5.3)
31	2006	8	5	29.865	80.188	14.5	5.1 (M_L 5.0)
32	2008	9	4	30.242	80.382	08.9	5.0
33	2009	10	3	30.018	79.827	24.5	5.1 (M_L 5.1)
34	2010	2	22	29.987	80.069	20.5	5.2 (M_L 5.3)
35	2010	6	22	29.910	80.460	20.4	5.1
36	2011	4	4	29.627	80.729	17.4	5.6
37	2011	5	4	30.216	80.415	23	5.1 (M_L 5.0)

Table 1: Contd.,							
38	2011	6	20	30.551	79.319	26.6	5.2 (M_L 5.2)
39	2012	2	26	29.614	81.005	17	5.1 (M_L 5.0)
40	2012	7	28	29.852	80.599	23	5.2 (M_L 5.2)
41	2012	11	11	29.405	81.447	27	5.4 (M_L 5.6)
42	2015	4	1	30.515	79.539	18	5.0
43	2016	12	1	29.902	80.551	21.3	5.2
44	2017	2	6	30.670	79.210	10	5.3

(*) For some of the earthquakes, the original values of other types of magnitude, as reported in the ISC catalog, are provided within brackets alongside the corresponding converted value.

x Not available.

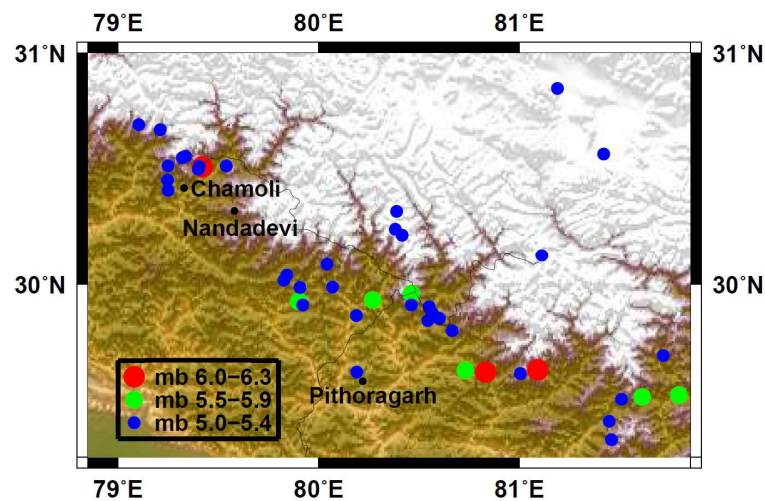


Figure 1: The Study Area in the Kumaun-Garhwal Region of Central Himalaya Showing Epicentral Locations (Colour-Coded Solid Circles) of Independent Earthquakes ($5 \leq m_b \leq 6.3$) that Occurred there During the Period 1958–2017 (As Listed in Table 1)

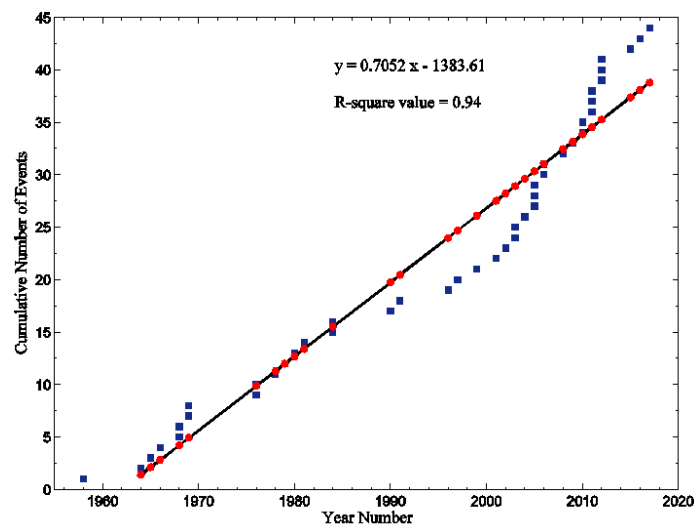


Figure 2: Illustration of the Time-Completeness Test Conforming to VCM (Mulargia and Tinti, 1985) Operated on the Catalog of Earthquakes in the Study Region that Gave the Working Data Set in the Magnitude Range $5 \leq m_b \leq 6.3$ (As Listed in Table 1) Covering a Period of Six Decades Since 1958. The Best-Fit Linear Trend through the Data Points of the Cumulative Number of Earthquakes as a Function of Time Confirms 'Completeness'

Table 2: Probability Distributions, Density Functions (PDF) and Associated Fisher Information Matrices (FIM)

Probability	Density Function		Parameters		FIM ($I(\theta)$)
Distribution	Pdf	Domain	Role	Domain	
Exponential	$\frac{1}{\alpha} e^{-\frac{t}{\alpha}}$	$t > 0$	α -scale	$\alpha > 0$	$\frac{1}{\alpha^2}$
Gamma †	$\frac{1}{\Gamma(\beta)} \frac{t^{\beta-1}}{\alpha^\beta} e^{-\frac{t}{\alpha}}$	$t > 0$	α -scale β -shape	$\alpha > 0$ $\beta > 0$	$\begin{bmatrix} \frac{\beta}{\alpha^2} & \frac{1}{\alpha} \\ \frac{1}{\alpha} & \psi'(\beta) \end{bmatrix}$
Lognormal	$\frac{1}{t\beta\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln t - \alpha}{\beta}\right)^2\right]$	$t > 0$	α -log-scale β -shape	$-\infty < \alpha < \infty$ $\beta > 0$	$\begin{bmatrix} \frac{2}{\beta^2} & 0 \\ 0 & \frac{1}{\beta^2} \end{bmatrix}$
Weibull †	$\frac{\beta}{\alpha^\beta} t^{\beta-1} e^{-\left(\frac{t}{\alpha}\right)^\beta}$	$t > 0$	α - β -scale shape	$\alpha > 0$ $\beta > 0$	$\begin{bmatrix} \frac{\beta^2}{\alpha^2} & -\frac{1}{\alpha}(1+\psi(1)) \\ -\frac{1}{\alpha}(1+\psi(1)) & \frac{1}{\beta^2}(\psi'(1)+\psi^2(2)) \end{bmatrix}$

† $\psi(x)$ and $\psi'(x)$ denote the digamma function and its first derivative respectively

Table 3: Estimated Values of the Scale (α) and Shape (β) Parameters Along with Their Asymptotic Standard Deviations and Confidence Bounds for Four Different Types of Probability Distribution Models Studied

Model	Parameter Value		Asymptotic Standard Deviation		Confidence Interval (95%)	
					Lower	Upper
Exponential	α	1.3522	σ_α	0.2062	0.9480	1.7564
Gamma	α	1.6396	σ_α	0.4116	0.8329	2.4463
	β	0.8248	σ_β	0.1539	0.5232	1.1264
Lognormal	α	-0.3044	σ_α	0.1227	-0.5449	-0.0639
	β	1.1381	σ_β	0.1736	0.7978	1.4784
Weibull	α	1.3004	σ_α	0.2250	0.8594	1.7414
	β	0.9281	σ_β	0.1104	0.7117	1.1445

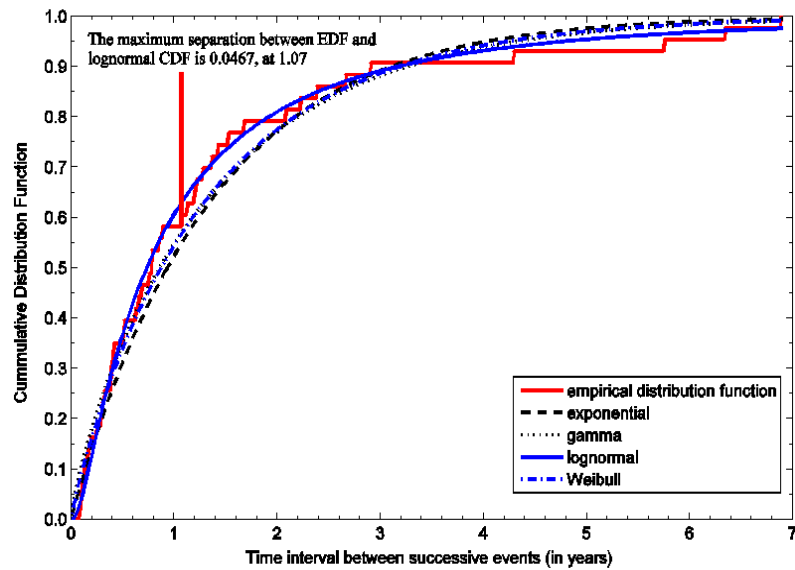


Figure 3: Illustrative Plot of the K-S Test Showing Cumulative Distribution Function as the Difference (Separation) between the Estimated Distribution Function and the Empirical Distribution Function for Each of the four Probability Models Applied to the Observed Earthquake Data Set

Table 4: Model Selection and Validation by two Different Goodness-of-Fit Tests

Model	Maximum Likelihood Test		K-S Test
distribution	$\ln L$	AIC	K-S distance
Exponential	-55.98	113.95	0.0992
Gamma	-56.25	116.49	0.0864
Lognormal	-53.49	110.97	0.0467
Weibull	-55.75	115.50	0.0800

$\ln L$: log-likelihood value; AIC : value using Akaike Information Criterion.

Table 5: Estimates, Based on Lognormal Distribution Model, of the Conditional Probabilities of an Earthquake ($5 \leq m_b \leq 6.3$) to Occur in the Next v Years in the Study Region, Assuming Absence of Earthquake in the Same Magnitude Range over τ Years (Elapsed Time) Since the Last Event on 06 February 2017

v (Years)	$\tau \rightarrow$ Elapsed Time (Years) Since the Last Listed Event on February 06, 2017				
	1 (2018)	2 (2019)	3 (2020)	4 (2021)	5 (2022)
1	0.52	0.43	0.37	0.33	0.29
2	0.72	0.64	0.57	0.52	0.48
3	0.83	0.76	0.70	0.65	0.61
4	0.88	0.83	0.78	0.74	0.70
5	0.92	0.87	0.83	0.80	0.76

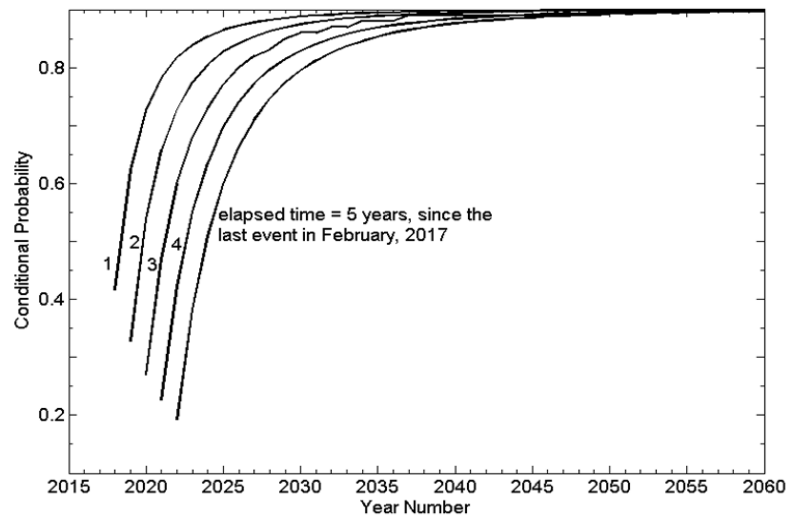


Figure 4: Conditional Probability Curves (Construed as Seismic Hazard Curves) of Occurrence of an Earthquake of Magnitude $5 \leq m_b \leq 6.3$ in the Coming Years in the Study Region for an Elapsed Time (Assumed Earthquake Absence) of $\tau = 1, 2, 3, 4, 5$ Years since the Last Event on 06 February 2017

